

KAJIAN TEORI

A. Penilaian.

Penilaian merupakan hal sangat penting dalam proses pembelajaran. Penilaian digunakan untuk mengambil keputusan penting terkait peserta didik, seperti menentukan apakah peserta didik tersebut perlu mengulang materi, naik kelas, mengulang atau tidak. Diperlukan pertimbangan yang matang untuk agar diperoleh keputusan yang tepat sehingga tidak merugikan peserta didik.

Untuk mendapatkan keputusan yang tepat, diperlukan informasi yang memadai tentang peserta didik, seperti penguasaan terhadap materi, sikap dan perilakunya. Dalam konteks ini penilaian memegang peranan yang cukup penting. Dari sini penilaian diharapkan memberi umpan balik yang objektif tentang apa yang telah dipelajari oleh peserta didik, bagaimana mereka belajar dan digunakan untuk mengetahui efektifitas dari proses pembelajaran.¹

Dengan demikian, apabila guru memiliki kemampuan untuk melakukan penilaian dengan baik maka dipastikan ia memiliki kemampuan mengajar yang baik pula. Uraian tersebut menandakan bahwa untuk menjadikan proses pembelajaran berkualitas maka guru seharusnya menguasai teknik penilaian yang baik pula. Sebab pembelajaran dan penilaian merupakan dua unsur yang tidak bisa dipisahkan dalam kegiatan belajar

¹ Kusaeri dan Suprananto, *Pengukuran dan Penilaian Pendidikan*, (Yogyakarta: Graha Ilmu, 2012) 8.

mengajar. Selanjutnya akan dijelaskan mengenai konsep dasar penilaian dan bagaimana cara memilih teknik penilaian yang tepat.

1. Konsep Dasar Penilaian.

Ada tiga istilah yang sering dipakai orang secara rancu, yaitu pengukuran, penilaian, dan evaluasi. Ketiga istilah ini memiliki arti yang sangat berbeda karena tingkat penggunaannya yang berbeda.²

Pengukuran merupakan cabang dari ilmu statistika terapan yang bertujuan untuk membangun dasar-dasar pengembangan tes sehingga dapat menghasilkan tes yang valid dan reliabel. Arikunto mendefinisikan bahwa mengukur adalah membandingkan sesuatu dengan satu ukuran, pengukuran bersifat kuantitatif.³ Azwar mendefinisikan pengukuran sebagai suatu prosedur pemberian angka terhadap atribut atau variabel sepanjang kontinum.⁴ Dengan demikian, secara sederhana pengukuran dapat dikatakan sebagai suatu prosedur membandingkan antara atribut yang hendak diukur dengan alat ukurnya.

Penilaian lazimnya dimulai dari pengukuran. Menurut Gronlund & Linn penilaian adalah suatu proses yang sistematis dan mencakup kegiatan mengumpulkan, menganalisa, serta menginterpretasikan informasi untuk menentukan seberapa jauh seseorang mencapai tujuan pembelajaran yang ditetapkan.⁵ Jadi penilaian adalah suatu proses yang

² Kusaeri, *Acuan dan Teknik Penilaian Proses dan Hasil Belajar dalam Kurikulum 2013*, (Yogyakarta : Ar-Ruzz Media, 2014) 14.

³ Suharsimi Arikunto, *Evaluasi Program Pendidikan*, (Jakarta : Bumi Aksara, 2008) 3.

⁴ Azwar, *Sikap Manusia: Teori Dan Pengukurannya*, (Yogyakarta, Pustaka Pelajar, 1995) 3.

⁵ Gronlund & Linn, *Measurement And Evaluation In Teaching*, (New York, Mac Millan Publishing, 1990) 5.

Evaluasi adalah kegiatan identifikasi untuk melihat apakah suatu program yang telah direncanakan telah tercapai atau belum, berharga atau tidak berharga, dan dapat pula untuk melihat tingkat efisiensi pelaksanaannya. Evaluasi adalah suatu proses penilaian untuk mengambil keputusan yang menggunakan seperangkat hasil pengukuran dan berpatokan kepada tujuan yang telah dirumuskan.⁶ Sehingga Kegiatan evaluasi dilakukan dengan cara membandingkan hasil pengukuran dan penilaian dengan suatu norma atau kriteria, dan hasilnya dinyatakan secara evaluatif.

2. Pemilihan Teknik Penilaian.

⁶ Dadan Rosana, *Modul Evaluasi UT Bab I*, (Yogyakarta, 2011) 12.

Pertama, tujuan pembelajaran (dalam konteks sekarang dalam bentuk kompetensi dasar dan dirinci sebagai indikator). Sebelum menilai peserta didik guru harus menentukan tujuan pembelajaran. Semakin jelas dan spesifik tujuan pembelajaran semakin mudah dalam menentukan teknik penilaian yang tepat.

Ketiga, teknik penilaian yang dipilih harus sesuai dengan karakteristik dan kebutuhan peserta didik. Pemilihan teknik penilaian yang tepat tidak hanya membantu peserta didik memperoleh informasi tentang hasil belajar namun juga akan sangat bermakna.

Keempat, dalam menginterpretasikan hasil penilaian guru harus mempertimbangkan kelemahan setiap teknik penilaian. Meskipun guru menggunakan teknik penilaian tertentu, informasi sebenarnya yang diperoleh adalah sebagian saja. Sehingga diperbolehkan menggunakan beberapa teknik penilaian untuk mengukur kemampuan peserta didik.⁷

[illegible]

Tes merupakan salah satu alat untuk melakukan pengukuran, yaitu alat untuk mengumpulkan informasi karakteristik suatu objek. Dalam pembelajaran, objek yang dimaksud adalah kecakapan peserta didik, minat, motivasi, dan sebagainya.⁸ Menurut Djemari Mardapi tes merupakan salah satu cara untuk menaksir besarnya kemampuan seseorang secara tidak langsung, yaitu dengan cara memberikan stimulus atau pertanyaan untuk mengetahui respon dari orang tersebut.⁹ Menurut Suharsimi Arikunto tes adalah suatu alat atau prosedur yang sistematis dan obyektif untuk memperoleh data-data atau keterangan yang diinginkan tentang seseorang, dengan cara yang boleh dikatakan tepat dan cepat.¹⁰ Menurut Kusaeri tes adalah suatu prosedur yang sistematis untuk menggambarkan karakteristik tertentu tentang peserta didik dengan menggunakan deskripsi dan angka.¹¹

Dari beberapa pendapat para ahli terkait pengertian tes, disimpulkan bahwa tes dalam kegiatan pembelajaran adalah serangkaian kegiatan yang

¹¹ Kusaeri, *Acuan dan Teknik Penilaian*, 14.

Menurut sejarah, tes yang pertama kali digunakan adalah tes tulis bentuk uraian. Tes tertulis bentuk uraian adalah Teori Tes Klasik atau *Classical True-Score Theory*, dinamakan Teori Tes Klasik karena unsur-unsur teori ini sudah dikembangkan dan diaplikasikan sejak lama, namun tetap bertahan hingga sekarang.¹³

Tes tertulis bentuk uraian merupakan seperangkat soal yang berupa tugas, pertanyaan yang menuntut peserta didik untuk mengorganisasikan dan menyatakan jawabannya menurut kata-kata sendiri. Jawaban tersebut dapat berbentuk mengingat kembali, menyusun, mengorganisasikan atau memadukan pengetahuan yang telah dipelajarinya dalam rangkaian kalimat

21. Sumardi Suryabrata, "*Pengembangan alat ukur psikologis*," (Yogyakarta: Andi Offset, 2002)

Eko Putro Widoyoko menambahkan bahwa penskoran tes uraian non objektif dipengaruhi oleh pemberi skor. Jawaban yang sama dapat memiliki skor yang berbeda oleh pemberi skor yang berbeda. Hal ini diakibatkan oleh beberapa hal, antara lain (a) Ketidak konsistenan penilai (b) *Hallo effect* atau kesan guru terhadap peserta didik sebelumnya (c) Pengaruh urutan pemeriksaan (d) Pengaruh bentuk tulisan dan bahasa.¹⁷

¹⁷ Eko Putro Widyoko, *Evaluasi*, 47.

Secara umum penulisan tes tertulis bentuk uraian non objektif harus memperhatikan beberapa hal sebagai berikut:¹⁸ (a) Mengukur kompetensi peserta didik. Artinya soal uraian tersebut mampu mengukur kemampuan peserta didik secara nyata dan akurat. (b) Soal uraian mampu mendorong peserta didik untuk berlogika dan berpikir tingkat tinggi. (c) Mengukur kemampuan berpikir kritis. (d) Materi yang diujikan hendaknya materi yang mampu merangsang kemampuan peserta didik untuk memecahkan masalah. (e) Pertanyaan yang diujikan hendaknya menggunakan kata kerja yang jelas dan mudah dipahami peserta didik. (f) Setiap soal harus mempunyai rubrik penskoran, dengan demikian hasil koreksi jawaban bisa lebih akurat.

¹⁸ Kunandar, *Penilaian Autentik*, 211.

[illegible]

(ruang lingkup) harus jelas. (3) Isi materi sesuai dengan tujuan pengukuran. (4) Isi materi yang ditanyakan sesuai dengan jenjang, jenis sekolah dan tingkat kelas.

Kedua, konstruksi²⁰ (1) Rumusan kalimat soal atau pertanyaan harus menggunakan kata tanya atau perintah yang menuntut jawaban terurai; seperti : mengapa, uraikan, jelaskan, hubungkan, tafsirkan, buktikan, hitunglah, dsb. Jangan menggunakan kata Tanya yang tidak menuntut jawaban uraian, misalnya: siapa, dimana, kapan. Demikian juga jangan menggunakan kalimat tanya yang menuntut jawaban ya atau tidak. (2) Buatlah petunjuk yang jelas tentang cara mengerjakan soal. (3) Buatlah pedoman penyekoran segera setelah soal selesai ditulis dengan cara menguraikan komponen yang akan dinilai atau kriteria penskorannya, besarnya skor bagi setiap komponen, serta rentang skor yang dapat diperoleh untuk soal yang bersangkutan. (4) Hal-hal lain yang menyertai soal (grafik, tabel, gambar, peta, atau yang sejenisnya) harus jelas dan terbaca, sehingga tidak menimbulkan penafsiran yang berbeda.

Ketiga, bahasa²¹ (1) Rumusan kalimat soal harus komunikatif, yaitu menggunakan bahasa yang sederhana, dan menggunakan kata-kata yang sudah dikenal siswa, serta baik dari segi kaidah bahasa Indonesia. (2) Butir soal menggunakan bahasa Indonesia yang baik dan benar. (3) Rumusan soal tidak menggunakan kata/kalimat yang menimbulkan

²⁰ Kusaeri, *Acuan Dan Teknik Penilaian*, 92.

²¹ *Ibid.*, 93.

Untuk mengoreksi tes tertulis bentuk uraian non objektif dapat dilakukan dengan menggunakan metode *point method* dan *rating method*.²² *Point method* adalah metode pengoreksian dengan cara membandingkan setiap jawaban dengan jawaban ideal yang telah ditetapkan dalam rubrik penskoran. Skor yang diberikan kepada setiap jawaban akan tergantung pada derajat kepadanannya dengan rubrik penskoran.

Rating method adalah metode pengoreksian dengan cara setiap jawaban siswa ditetapkan dalam salah satu kelompok yang sudah dipilah-pilah berdasarkan mutunya selagi jawaban tersebut di baca. Kelompok-kelompok tersebut menyatakan mutu dan menentukan berapa skor yang dapat diberikan kepada setiap jawaban. Misalnya sebuah soal akan diberi skor maksimum 8, maka bagi soal tersebut dapat dibuat 9 kelompok jawaban dari 8 sampai 0.

Djemari Mardapi menambahkan bahwa untuk mengoreksi soal uraian hendaknya dilakukan dengan cara menilai jawaban pertanyaan demi pertanyaan bukan peserta didik ke peserta didik. Selanjutnya seorang guru menghilangkan identitas peserta didik dan menggantinya

[illegible]

Sedangkan menurut Kunandar, ada beberapa langkah untuk mengoreksi soal bentuk uraian non objektif agar mendekati objektif yaitu:²⁴ (a) menyusun pola jawaban yang diambil dari sampel jawaban peserta didik (b) pemeriksaan jawaban tidak dilakukan dengan cara membaca jawaban satu peserta didik namun dengan cara pernomor (c) setiap lembar jawaban dikoreksi lebih dari satu kali (d) nilai peserta didik tidak langsung dijumlahkan secara global tetapi dirinci setiap aspek-aspek penilaian.

Dalam penelitian ini digunakan metode pengoreksian *point method*, dengan beberapa tambahan dari Djemari Mardapi. Hal ini bertujuan untuk meningkatkan keakuratan dari pengkoreksian yang akan dilakukan.

3. Penyusunan Pedoman Penskoran.

Pedoman penskoran merupakan panduan atau petunjuk yang menjelaskan tentang: Batasan atau kata-kata kunci untuk melakukan penyekoran terhadap soal-soal bentuk uraian dan kriteria-kriteria jawaban yang digunakan untuk melakukan penyekoran terhadap soal-soal bentuk uraian non-objektif. Dengan pedoman atau rubrik penskoran, guru dapat mengoreksi jawaban peserta didik secara akurat. Pedoman penskoran

²⁴ Kunandar, *Penilaian Autentik*, 246.

hendaknya disusun segera setelah perumusan kalimat butir-butir soal untuk menjaga keobjektivitasan dari penilaian yang akan dilakukan.²⁵

Sedangkan kelemahan dari Tes tertulis bentuk uraian non objektif adalah sebagai berikut: (a) Sampel soal sangat terbatas sehingga bahan materi yang diujikan juga terbatas. (b) Cara memeriksa hasil tes sulit dan bisa mengandung unsur subyektivitas. (c) Membutuhkan waktu yang lama untuk proses koreksi. (d) Membutuhkan waktu yang lama untuk menyelesaikan satu soal uraian. (e) Tidak banyak kompetensi dasar yang dapat diuji.²⁹

²⁹ *Ibid.*, 214.

Kunandar juga menyatakan bahwa tes tertulis bentuk uraian non objektif dapat digunakan untuk mengukur kemampuan hasil belajar peserta didik pada tingkat C4, C5, C6. Karena tes tertulis bentuk uraian non objektif dapat menilai berbagai jenis kemampuan seperti: mengemukakan pendapat, berpikir kritis, berpikir kreatif dan pemecahan masalah.³³

Oleh karena itu, dalam merumuskan butir soal untuk tes tertulis bentuk uraian non objektif harus memperhatikan kemampuan peserta didik pada tingkat menganalisa, mengevaluasi dan mencipta.

Untuk melakukan penilaian yang baik maka perumusannya tidak bisa dipisahkan dari tujuan pembelajaran. Penilaian yang baik diturunkan dari tujuan pembelajaran yang jelas. Tujuan pembelajaran yang jelas akan sangat membantu agar penilaian yang dilakukan benar-benar mengukur apa yang telah diajarkan kepada peserta didik.³⁴

³² Masidjo, *Penilaian Pencapaian Hasil Belajar Siswa Disekolah*, (Kanisius: Yogyakarta, 1995) 46.

³³ Kunandar, *Penilaian Autentik*, 209

³⁴ Kusaeri, *Acuan Dan Teknik Penilaian*, 30.

(perilaku yang harus ditampilkan), *condition* (kondisi yang diberikan), dan *degree* (tingkatan yang diberikan).³⁵

Para ahli kurikulum telah sepakat untuk melakukan klasifikasi (taksonomi) tujuan pembelajaran. Terdapat bermacam-macam taksonomi tujuan pembelajaran, taksonomi tersebut diberi nama sesuai dengan nama penciptanya. Salah satu rujukan dalam sistem pendidikan nasional untuk merumuskan tujuan pembelajaran adalah Taksonomi Bloom.³⁶

Taksonomi pada dasarnya merupakan usaha pengelompokan yang disusun dan diurut berdasarkan ciri-ciri tertentu. Menurut Dadan Rosana taksonomi tujuan pembelajaran sangat diperlukan, karena pertimbangan sebagai berikut: (a) Perlu adanya kejelasan terminologi yang digunakan dalam tujuan pembelajaran, sebab tujuan pembelajaran berfungsi untuk memberikan arah kepada proses belajar dan menentukan perilaku yang dianggap sebagai bukti belajar. (b) Sebagai alat yang akan membantu guru dalam mendeskripsikan dan menyusun tes, teknik penilaian dan evaluasi.³⁷

Berdasarkan pertimbangan tersebut, diketahui bahwa taksonomi tujuan pembelajaran dapat membantu guru dalam penyusunan tes. Oleh karena itu, dalam proses penyusunan tes tertulis bentuk uraian non objektif untuk pembelajaran berbasis masalah yang baik, hendaknya

³⁵ Depdiknas, *Kurikulum Tingkat Satuan Pendidikan*, (Jakarta: Balitbang Depdiknas 2009) 14.

³⁶ Kusaeri, *Acuan Dan Teknik Penilaian*, 33.

³⁷ Dadan Rosana, *Modul Evaluasi UT Bab I*, (Yogyakarta, 2011) 46.

mengacu pada salah satu model taksonomi tujuan pembelajaran yang dikemukakan oleh para ahli.

Tes tertulis bentuk uraian non objektif merupakan tes yang bertujuan untuk mengukur kemampuan kognitif peserta didik. Maka dalam penelitian ini menggunakan Taksonomi Bloom edisi revisi domain kognitif. Taksonomi Bloom revisi dimensi proses kognitif yang berisikan enam kategori pokok, dengan jenjang yang paling rendah sampai jenjang yang paling tinggi. Selain domain kognitif dalam perumusannya juga memperhatikan dimensi pengetahuan yang meliputi pengetahuan faktual, konseptual, prosedural dan metakognitif.

Tingkatan berpikir Taksonomi Bloom edisi revisi adalah sebagai berikut: (a) Mengingat (*remember*) yaitu mengingat kembali pengetahuan yang relevan dari memori jangka panjang. (b) Memahami (*understand*) yaitu membangun pengetahuan dari pesan pembelajaran, termasuk komunikasi lisan, tertulis, dan grafis. (c) Menerapkan (*apply*) yaitu melaksanakan atau menggunakan prosedur dalam suatu situasi tertentu. (d) Menganalisis (*analyze*) yaitu memecah materi ke dalam bagian-bagian penyusunannya, dan menentukan bagaimana bagian-bagian tersebut saling berhubungan satu sama lain. (e) Mengevaluasi (*evaluate*) yaitu melakukan penilaian berdasarkan kriteria dan standar tertentu. (f) Menciptakan (*create*) yaitu menempatkan beberapa elemen secara bersamasama untuk membangun suatu keseluruhan yang logis dan

fungsional, dan mengatur elemen-elemen tersebut ke dalam pola atau struktur yang baru.³⁸

Kemampuan pada tingkatan menganalisis, mengevaluasi, dan mengkreasi merupakan kemampuan berpikir tingkat tinggi, sedangkan kemampuan pada tingkatan mengingat, memahami, dan menerapkan merupakan kemampuan tingkat rendah.³⁹ Pembelajaran berbasis masalah merupakan suatu model pembelajaran untuk melatihkan kemampuan berpikir tingkat tinggi peserta didik.⁴⁰ Oleh karena itu tes tertulis bentuk uraian non objektif yang disusun, hendaknya mengacu pada kemampuan berpikir tingkat tinggi pada level menganalisis, mengevaluasi, dan mengkreasi.

Kemampuan berpikir tingkat tinggi dibagi menjadi dua, yaitu berpikir kritis dan berpikir kreatif. Berpikir kritis merupakan kemampuan memberikan rasionalisasi terhadap sesuatu dan mampu memberikan penilaian terhadap sesuatu tersebut. Sedangkan berpikir kreatif adalah kemampuan untuk melakukan generalisasi dengan menggabungkan, mengubah atau mengulang kembali keberadaan ide-ide tersebut.⁴¹

Kemampuan peserta didik dalam berpikir kritis dan kreatif perlu

³⁸ Lorin Anderson and Krathwohl, *A taxonomy for Learning, Teaching, and Assessing*, (New York : Addison Wesley Longman, Inc, 2001) 67.

³⁹ Rini Julistiawati, *Keterampilan Berpikir Level C4, C5, & C6 Revisi Taksonomi Bloom Siswa Kelas X-3 Sman 1 Sumenep Pada Penerapan Model Pembelajaran Inkuiri*, *Journal of Chemical Education* Vol. 2 No.2 (Mei, 2013), 58.

⁴⁰ Wina Sanjaya, *Strategi Pembelajaran (Berorientasi Standar Proses Pendidikan)*, (Jakarta: Kencana, 2007) 218.

⁴¹ Kunandar, *Penilaian Autentik*, 171.

Dalam taksonomi bloom revisi juga diuraikan tentang klasifikasi dimensi pengetahuan dalam empat kategori, yaitu pengetahuan faktual, konseptual, prosedural dan metakognitif.⁴² Pengetahuan faktual berisikan pengetahuan tentang elemen dasar yang harus diketahui siswa untuk mengenal satu disiplin ilmu atau untuk menyelesaikan masalah didalamnya. Pengetahuan ini meliputi Pengetahuan tentang istilah dan pengetahuan tentang rincian dan unsur tertentu.

Pengetahuan Prosedural yaitu pengetahuan tentang bagaimana melakukan suatu hal, metode dan inquiri, dan kriteria untuk menggunakan suatu keterampilan, algoritma, teknik dan suatu metode. Pengetahuan ini mencakup Pengetahuan tentang keterampilan dan algoritma tertentu, Pengetahuan tentang teknik dan metode tertentu dan Pengetahuan tentang kriteria untuk menentukan kapan menggunakan prosedur yang tepat.

[illegible]

2. Penyusunan Tes Tertulis Bentuk Uraian Non Objektif Untuk Pembelajaran Berbasis Masalah.

Untuk menjamin keakuratan soal tes tertulis bentuk uraian non objektif, maka soal harus memenuhi syarat-syarat sebagai berikut: (a) membatasi ruang lingkup dengan memilih materi atau bahan pelajaran yang esensial (b) menggunakan bahasa yang baik dan benar sehingga mudah difahami dengan baik oleh peserta didik (c) jangan mengulang pertanyaan pada materi yang sama (d) tuliskan rubrik penskoran sebelum

D. Analisis Kualitas Soal.

⁴³ Kunandar, *Penilaian Autentik*, 212.

⁴⁴ Nana Sudjana, *Penilaian Hasil Proses Belajar Mengajar*, (Bandung: Remaja Rosdakarya, 2013) 135.

⁴⁵ Zainal Arifin, *Evaluasi Pembelajaran*, 246.

Suatu tes dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya atau memberikan hasil ukur yang tepat sesuai dengan apa yang hendak diukur. Sumarna Supranata berpendapat bahwa “Validitas merupakan suatu konsep yang berkaitan dengan sejauh mana suatu tes dapat mengukur apa yang hendak diukur”. Validitas tes, secara keseluruhan ada empat macam validitas, yaitu: validitas isi (*content validity*), validitas konstruk (*construct validity*), validitas prediktif (*predictive validity*), dan validitas bandingan (*concurrent validity*).⁴⁸

[illegible]

Validitas konstruk menunjuk sejauh mana tes dapat mengukur dengan tepat aspek berpikir yang telah ditentukan dalam tujuan instruksional secara khusus.⁵⁰ Validitas konstruk dapat dilakukan dengan cara mencocokkan aspek-aspek berpikir dalam tes dengan aspek berpikir yang dikehendaki dalam tujuan intruksional khusus. Dalam hal ini, pengerjaannya didasarkan pada logika. Selain itu, dapat juga dilakukan dengan cara melakukan diskusi dengan orang yang ahli di bidang yang bersangkutan.⁵¹

Validitas bandingan menunjuk pada berapa jauh tes dapat mengukur tingkat penguasaan materi yang memang seharusnya dikuasai. Tes dikatakan memiliki validitas bandingan apabila tes tersebut dalam

⁵² Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Raja Grafindo Persada, 2011) 170.

waktu yang sama menunjukkan hubungan searah antara tes pertama dengan tes berikutnya.⁵³

Validitas yang digunakan dalam penelitian ini adalah validitas logis yang meliputi validitas isi dan validitas konstruk. Validitas tersebut diperoleh dengan cara penilaian para ahli melalui proses validasi.

2. Reliabilitas.

Reliabilitas disebut juga tingkat atau derajat konsistensi suatu tes. Tes akan dikatakan reliabel apabila diperoleh hasil yang sama ketika suatu instrumen diteskan pada kelompok yang sama di waktu yang berbeda.⁵⁴

Tinggi rendahnya koefisien reliabilitas sangat dipengaruhi oleh beberapa faktor. Crocker dan Algina menyebutkan bahwa faktor itu antara lain adalah panjang suatu tes, kecepatan, homogenitas belahan, dan tingkat kesukaran soal. Tingkat kesukaran soal memegang peranan yang paling dominan.

Untuk mengetahui reliabilitas suatu tes bisa menggunakan mekanisme: teknik *test-retest*, belah dua, dan bentuk ekuivalen.⁵⁵ Sedangkan menurut Sumarna Surapranata ada empat konsep reliabilitas yaitu: paralel atau ekuivalen, *test retest*, belah dua, dan *internal consistency*. Namun sebagian ahli berpendapat bahwa metode belah dua

⁵³ *Ibid.*,. 177.

⁵⁴ Zainal Arifin, *Evaluasi Pembelajaran*, 258.

⁵⁵ Djaali dan Puji Muljono, *Pengukuran*,. 57.

Bentuk tes ekuivalen merupakan dua buah tes yang dibuat setara seperti memiliki kesamaan tujuan, tingkat kesukaran dan susunan butir soal yang berbeda. Skor dari kedua kelompok tes dikorelasikan untuk mendapatkan reliabilitas soal.

Penelitian ini menggunakan teknik belah dua. Kenyataannya, terdapat berbagai cara untuk membelah tes. Untuk perangkat tes dengan jumlah soal sebanyak enam, terdapat sepuluh cara dan sepuluh

[illegible]

dibagi menjadi 27% kelompok atas, yaitu kelompok yang memiliki skor total tinggi dan 27% kelompok bawah, yaitu kelompok yang memiliki skor rendah.⁶² Menurut Suharsimi Arikunto Butir soal yang baik adalah butir soal yang mempunyai indeks diskriminasi 0,3 sampai dengan 0,70.⁶³